

Aprendendo com exemplos

1. Explique os princípios da Aprendizagem Supervisionada, Aprendizagem Não-supervisionada e Aprendizagem por reforço
2. Qual a diferença entre um classificador e um regressor? Exemplifique sua resposta para o problema de reconhecimento de dígitos.
3. A aprendizagem não supervisionada pode determinar sozinha se uma situação é boa ou ruim para um agente? Por quê?
4. Explique qual é o objetivo da aprendizagem indutiva. Do que precisamos para “aprender” uma hipótese h ?
5. Explique rapidamente como funciona o algoritmo de aprendizagem para árvores de decisão visto em sala de aula.
6. O que significa dizer que um atributo é inútil? E que um atributo é perfeito?
7. Foram recolhidos os dados abaixo (dados de treinamento) de uma linha de montagem com vários sensores. Pretende-se usar aprendizagem por árvores de decisão para identificar quais as peças a rejeitar por serem demasiado frágeis.

Amostra	Cor	Eco	Temperatura	Frágil
1	1	Forte	Alta	Sim
2	2	Forte	Normal	Sim
3	2	Fraco	Alta	Sim
4	3	Forte	Normal	Não

- a. Qual é a entropia do conjunto de exemplos acima? O que representa essa informação?
 - b. Que atributo será selecionado como primeiro atributo a dividir o conjunto? Por quê?
 - c. Desenhe a árvore de decisão correspondente (construída de acordo com o critério de maximização do ganho de informação). Justifique o resultado obtido.
8. Considere o problema de aprender o conceito de se ou não comprar um CD de música (poderíamos pensar em um sistema de recomendação de compras). Para simplificá-lo, tal que possamos resolvê-lo manualmente, usaremos um número muito pequeno de exemplos a partir do qual aprenderemos o conceito.

Suponha que você está usando os seguintes atributos para descrever os exemplos:

TIPO valores possíveis: (R)ock, (J)azz, (H)ipHop

PREÇO valores possíveis: (B)arato, (C)aro

O valor de saída é binário, assim usaremos ' + ' e ' - ' indicando uma recomendação da "compra" ou não, respectivamente.

O conjunto de treinamento é o seguinte:

TIPO = H PREÇO = C CATEGORIA = +
TIPO = R PREÇO = B CATEGORIA = +

TIPO = R PREÇO = C CATEGORIA = +
TIPO = H PREÇO = B CATEGORIA = +
TIPO = J PREÇO = B CATEGORIA = +
TIPO = R PREÇO = C CATEGORIA = -
TIPO = J PREÇO = C CATEGORIA = -
TIPO = J PREÇO = B CATEGORIA = -
TIPO = H PREÇO = C CATEGORIA = +
TIPO = J PREÇO = C CATEGORIA = -
TIPO = R PREÇO = C CATEGORIA = -
TIPO = J PREÇO = B CATEGORIA = +
TIPO = R PREÇO = C CATEGORIA = -

E o conjunto de teste:

TIPO = R PREÇO = B CATEGORIA = +
TIPO = J PREÇO = B CATEGORIA = -
TIPO = J PREÇO = C CATEGORIA = -
TIPO = R PREÇO = C CATEGORIA = +
TIPO = H PREÇO = C CATEGORIA = +

- a) Construção da árvore: aplique o algoritmo de árvore de decisão (com ganho de informação) ao conjunto de treinamento acima. Mostre todos os passos. Se há empate entre os valores do ganho de informação de vários atributos (na escolha do melhor deles), escolha aquele cujo nome apareça primeiro em ordem alfabética.
- b) Análise de desempenho: use a árvore de decisão produzida em (a) para classificar os exemplos do conjunto de teste. Informe a acurácia da árvore para esses exemplos (por exemplo, porcentagem de exemplos classificados corretamente). Discuta resumidamente os resultados.